

II Ciclo: DIÁLOGOS EN EL TERRITORIO

"Ciencia abierta: la importancia de los datos abiertos para la I+D"

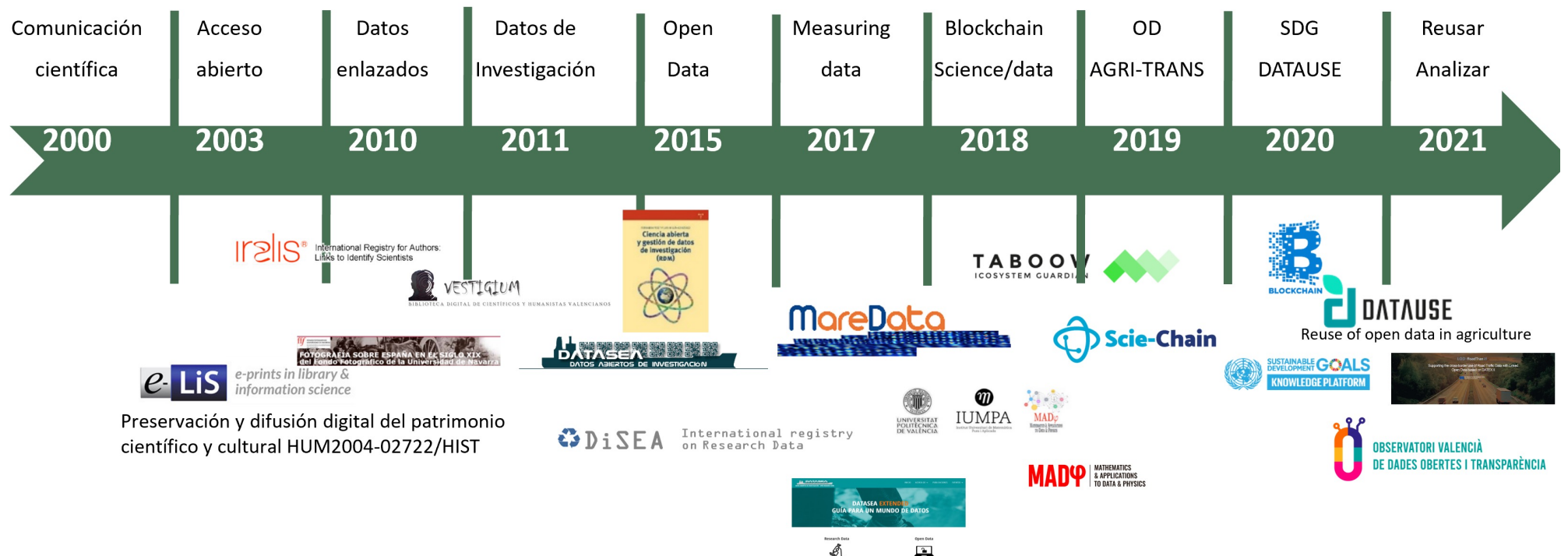


A cargo de:

Fernanda Peset. Catedrática de la Universitat Politècnica de València y
Antonia Ferrer. Profesora Titular de la Universitat Politècnica de València
y Directora de la Cátedra de Transparencia y Participación.

Citar como: Ferrer Sapena, A.; Peset, F. (2021) Ciencia abierta: la importancia de los datos abiertos en la I+D. En II Ciclo Diálogos en el Territorio. Universitat d'Alacant, 26/03/2021, 19 h.

TRAYECTORIA Y EQUIPO



✓ **Ciencia abierta: definiciones y contexto**

- ☐ La importancia de abrir datos
- ☐ Beneficios de gestionar bien
- ☐ ¿Qué se pide en los proyectos y en las revistas?
- ☐ ¿Cuál es la realidad actual?
- ☐ Conclusiones y debate

EL INVESTIGADOR, CENTRO DEL SISTEMA



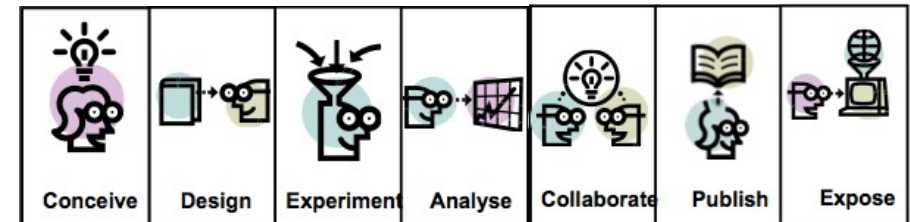
IUMPA
Institut Universitari de Matemàtica
Pura y Aplicada



Producción
Explotación
Networking/Data
sharing

Reconocimiento
Financiación
Rapidez

Preservación
Compartir



Su trabajo

Sus intereses

Nuevas
obligaciones y
openscience

CIENCIA ABIERTA. DEFINICIÓN



IUMPA
Institut Universitari de Matemàtica
Pura y Aplicada



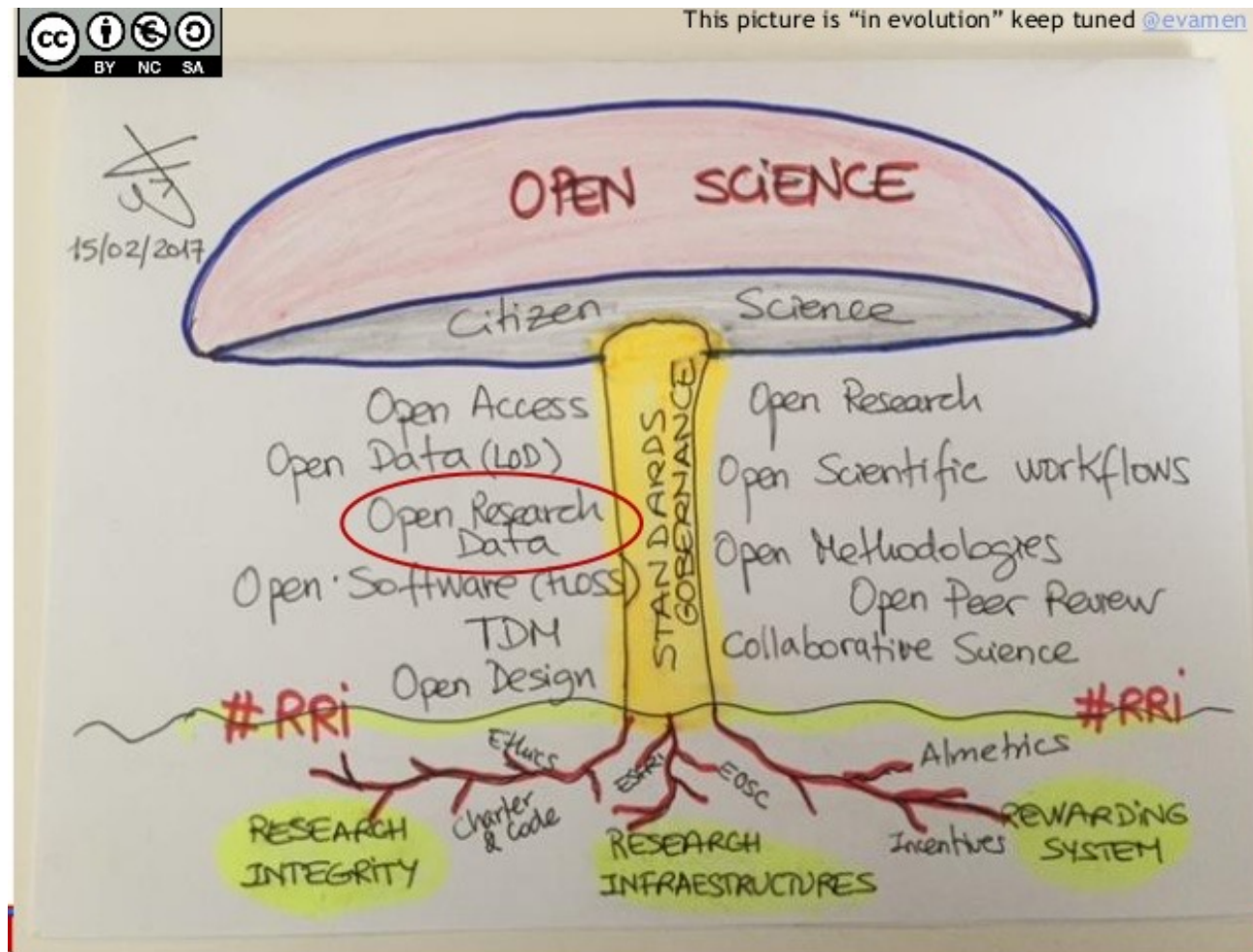
La ciencia abierta o Open Science es un movimiento cuyo objetivo es la accesibilidad de las investigaciones científicas para todos los ciudadanos. Este novedoso tipo de practicar la ciencia incluye la Investigación Abierta y el [Open notebook science](#), métodos de investigación en los que los experimentos y los datos obtenidos mediante éstos son automáticamente de acceso público, sobre todo al conocimiento científico producido con fondos públicos.)

WIKIPEDIA

Open Science is a system change allowing for better science through open and collaborative ways of producing and sharing knowledge and data, as early as possible in the research process, and for communicating and sharing results. This new approach affects research institutions and science practices by bringing about new ways of funding, evaluating and rewarding researchers. Open Science increases the quality and impact of science by fostering reproducibility and interdisciplinarity. It makes science more efficient through better sharing of resources, more reliable through better verification and more responsive to society's needs

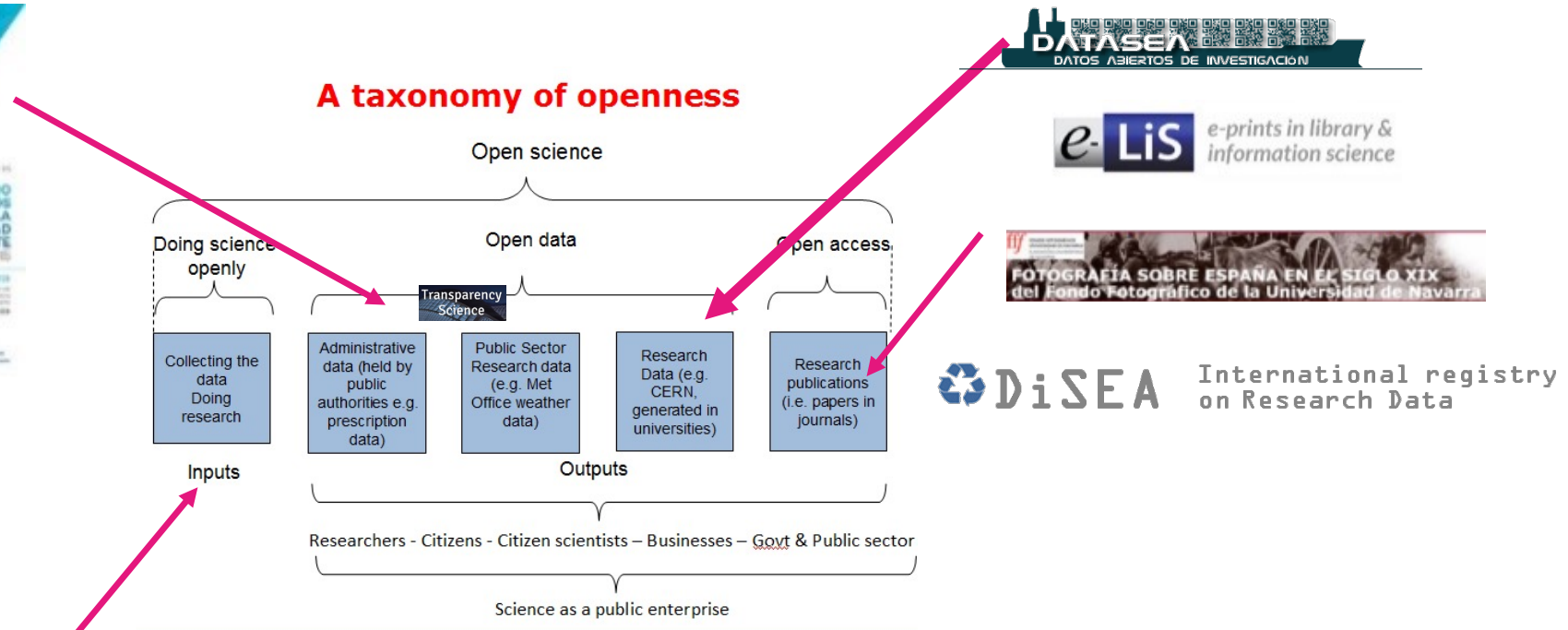
EUROPEAN COMMISSION

CIENCIA ABIERTA. COMPONENTES



Méndez, E. (2017). Open science

Los datos de investigación son parte de lo abierto!!! Directiva EU de reutilización 2019



A taxonomy of openness. Boulton, 2013

Open Science *does not equal* Open Access

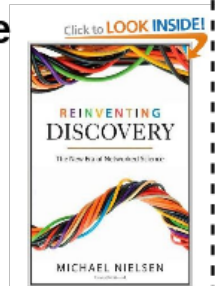
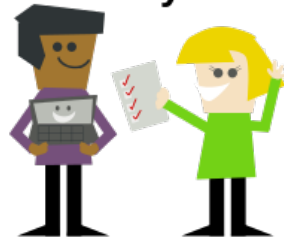
Open Data



Transparent
Processing



Community science



Engaging all stakeholders in the process
(data collection and analysis) and
synthesis of publically funded research.
*A 21st Century approach to engagement
a discussion*



Green Open Access

Pre-print or other version of a
publication held in a national or
institutional repository.



Gold Open Access

Payment to Open Access journal
to publish paper. Payment to
traditional journal to not place
paper behind a pay-wall. Held by
the journal.

Making the synthetic derivatives
of publically funded academic
research publically accessible.
*A 17th Century approach to engagement
a lecture*

Arbeck (2013).
http://commons.wikimedia.org/wiki/File:Open_Science_Does_Not_Equal_Open_Access.svg

Directiva 2019 sobre reutilización https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=uriserv:OJ.L_.2019.172.01.0056.01.SPA

Artículo 10 Datos de investigación

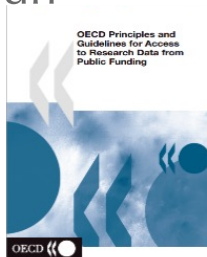
DATOS DE INVESTIGACIÓN. DEFINICIÓN Y EVIDENCIA



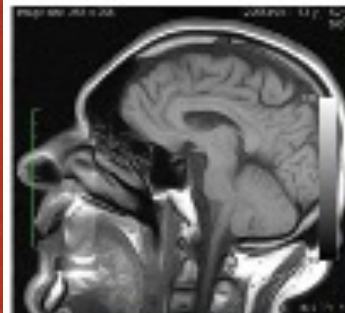
“datos de investigación [es] todo aquel material que ha sido registrado durante la investigación, reconocido por la comunidad científica y que sirve para certificar los resultados de la investigación que se realiza” (NIH) Torres–Salinas; Robinson–García; Cabezas–Clavijo, 2012

“...factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as **necessary to validate research findings**. A research data set constitutes a systematic, partial representation of the subject being investigated.”

Evaluados/certificados (*asessable*): asegurando que se ha realizado una evaluación sobre la credibilidad de los datos. Los datos han de proporcionar una suma de los resultados de un trabajo científico que es inteligible a quienes desean entenderlo o examinarlo.



Materia	Método del artículo	Material suplementario/datos que lo sustentan
Historia	Descriptivo...	Fuentes de archivo, no sujetas a derechos
Pedagogía	Encuesta	Cuestionarios, métodos, Entrevistas convenientemente anonimizadas
Literatura	Descriptivo	Obras de autores sin derechos vigentes, Voz
Informática	Programación	Código
Arte	Analítico/Descriptivo/Reflexivo	Descripciones de las obras, Comparaciones, objetos digitales





SENIOR MANGEMENT / RESEARCH FUNDERS

- provide curation infrastructure, guidance and support services
- policies / impetus for curation

RESEARCHERS

- create research data
- add context / meaning to data
- undertake early data management
- provide access to data
- reuse data

DISPOSE

REPOSITORY / DATA LIBRARIANS / ARCHIVISTS

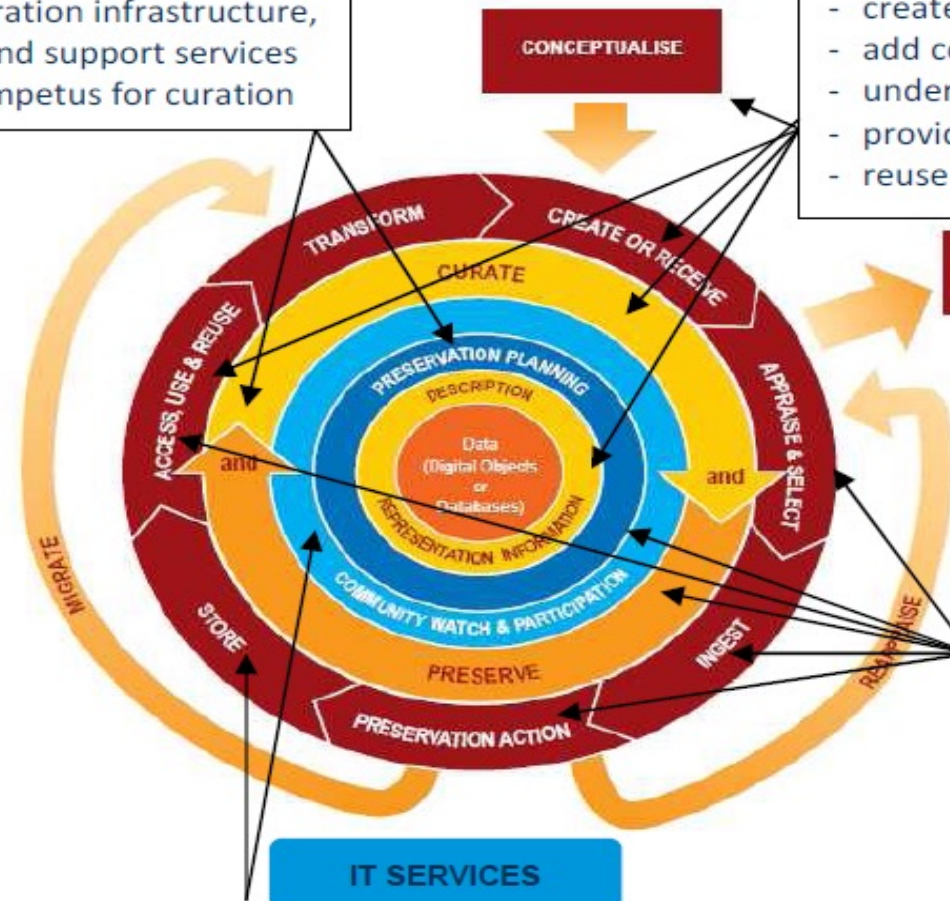
- help researchers select data
- ingest data to repository
- curate and preserve data
- facilitate access and reuse (including metadata creation)

IT SERVICES

- provide storage
- offer technical support

There are many stakeholders in the curation lifecycle with overlapping roles and responsibilities. Good data management requires collaboration across these groups.

To understand current practices and identify service gaps or areas for improvement, data surveys should engage with the whole range of people involved.



LIFE CYCLE MODEL (DCC)

✓ **Ciencia abierta: definiciones y contexto**

✓ **La importancia de abrir datos**

☐ Beneficios de gestionar bien

☐ ¿Qué se pide en los proyectos y en las revistas?

☐ ¿Cuál es la realidad actual?

☐ Conclusiones y debate

ESTRATEGIA EUROPEA DE DATOS



Publicada por la Comisión para hacer frente y poder competir en igualdad de oportunidades con otros países/continentes.

Europa tiene que estar preparada para un nuevo marco donde los productores de datos tengan un mayor control sobre los datos que generan.

EL VALOR DE LOS DATOS

Reside en su uso y reutilización.

Una de las ventajas que representan los datos es que se pueden **replicar a un coste casi nulo**, y además pueden utilizarse simultáneamente por distintas personas o instituciones

PROBLEMAS

- **No hay suficientes datos disponibles** para una reutilización innovadora, en particular en el caso del desarrollo de la inteligencia artificial. Datos tanto públicos como privados para intercambio entre empresas y administración, o a la inversa, o entre ambas.
- Existen **desequilibrios** en el poder de mercado. Son unos pocos los que lo dominan.
- La **interoperabilidad y calidad de los datos**, su estructura, autenticidad e integridad son clave para su explotación, especialmente en el contexto de la inteligencia artificial.

COMO TIENEN QUE SER

Por datos abiertos se entiende los datos en formatos abiertos que puede utilizar, reutilizar y compartir libremente cualquier persona con cualquier fin.

- Con **disponibilidad y reutilización generalizada** con fines privados o comerciales, con restricciones mínimas o nulas de carácter jurídico, técnico o económico.

«abiertos desde el diseño y por defecto»

- Deben garantizar la **protección de los objetivos de interés público**: la seguridad pública, la información sensible sobre infraestructuras críticas, garantizar la protección de los datos personales.

PRINCIPIOS FAIR

Se adaptaran a los principios FAIR.

Se **facilitará** también la **información sobre qué datos pueden ser utilizados, cómo y por quién**, cuando el fin sea la investigación científica.

Se proporcionará un mayor volumen de datos de calidad del sector público para su reutilización, teniendo especialmente en cuenta el potencial que puede tener para las pyme.



PRINCIPIOS FAIR



Enfatizan la capacidad de acción de las máquinas para: encontrar, acceder, interpretar y reutilizar.

Los principios se refieren a tres tipos de entidades: **datos**, **metadatos** e **infraestructura**.

¿Cuál es el coste de no cumplir los requisitos FAIR?

European Commission. Directorate General for Research and Innovation. PwC EU Services. (2019). Cost of not having FAIR research data. Cost-benefit analysis of FAIR research data. Retrieved from <https://doi.org/10.2777/02999>

EL COSTE



- **Consecuencia directa:**

- El coste anual de no tener datos de investigación FAIR supone a la economía europea al menos 10.200 millones de euros cada año.

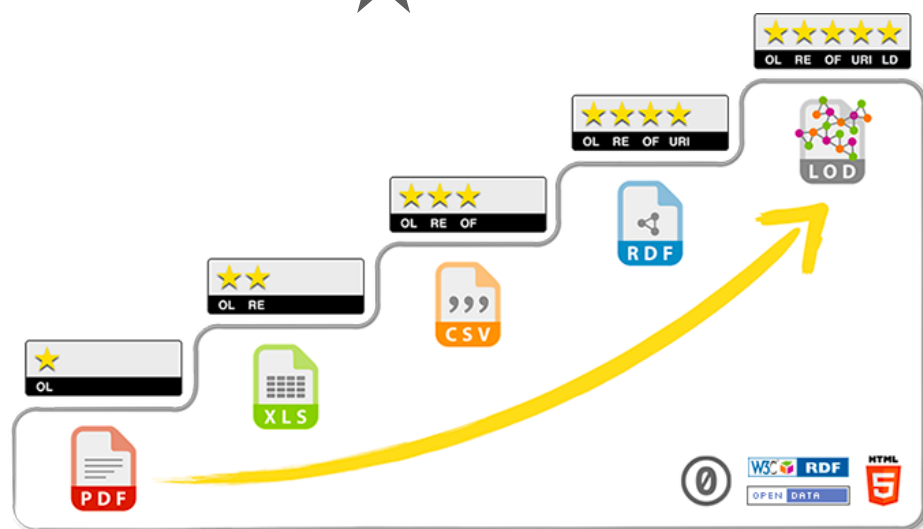
- **Indirectamente:**

- Impacto en la calidad de la investigación, la facturación económica o la legibilidad de los datos de investigación por parte de las máquinas.
- Puede suponer otros 16.000 millones de euros anuales además de lo estimado

European Commission. Directorate General for Research and Innovation. PwC EU Services. (2019). Cost of not having FAIR research data. Cost-benefit analysis of FAIR research data. Retrieved from <https://doi.org/10.2777/02999>

DATOS ABIERTOS – PRINCIPIOS FAIR

Modelo 5 ★ Datos Abiertos



Principio FAIR - no acumulativos

- Describe un proceso legible por máquinas o personas para acceder a los datos
- Describe el contexto en el que se generaron esos datos, para permitir la evaluación de su utilidad
- Define explícitamente las condiciones en las que se pueden reutilizar;
- Proporciona instrucciones sobre cómo deben citarse cuando se reutilizan.

Mons, B., Neylon, C., Velterop, J., Dumontier, M., da Silva Santos, L., & Wilkinson, M. (2017). Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. *Information Services & Use*, vol. 37, no. 1, pp. 49–56.

- ✓ **Ciencia abierta: definiciones y contexto**
- ✓ **La importancia de abrir datos**
- ✓ **Beneficios de gestionar bien**
- ☐ ¿Qué se pide en los proyectos y en las revistas?
- ☐ ¿Cuál es la realidad actual?
- ☐ Conclusiones y debate

BENEFICIOS DE GESTIONAR Y COMPARTIR LOS DATOS CORRECTAMENTE

DATASEA survey

- Aumentan las posibilidades de trabajar con grandes volúmenes de datos y de hacer análisis ciegos.
- Evita repetir costosos proyectos

Según el MIT

- incrementa el impacto de la investigación,
- ahorra tiempo y recursos,
- preserva a largo plazo los datos y asegura su integridad,
- atrae financiación externa,
- promueve nuevos descubrimientos y apoya lo abierto.

Para la ciencia y la sociedad en general (Boulton, 2013)

- permite responder de forma eficiente a las emergencias mundiales (pandemias o cambio climático);
- asegura la validación y reproducción de la ciencia;
- estimula la cooperación e innovación y descubrimientos adicionales;
- y en definitiva es un ejercicio de transparencia ante los ciudadanos que la financian con sus impuestos.

DATOS DE INVESTIGACIÓN. CÓMO GESTIONAR BIEN LOS DATOS

Electronic Laboratory Notebooks

Cumplen con requerimientos: almacenamiento, gestión del ciclo de vida, analítica y la preservación de los objetos de investigación.



Blockchain (LABII)

Comerciales (labGURU) Experimentales (Rohub)

Funcionales (Benchling)

¿QUÉ ES GESTIONAR BIEN?

Wigham (2014)

		Name	Example
Acadèmic	Digital_maps	Orig_image	Month_year_type
		Proc_image	Month_year_type_process
		Field_points	Field_points_date of actualization
		Mod_outcome_maps	Month_year_outcome type
		Scenario_maps	Year_parameter
	Field_data	(Parameter name)	Parameter_date of actualization
	Modelling	Orig_scripts	Model_script_version
		Modified_scripts	Model_script_modification_trial#
		Input_data	Parameter_col#_row#_trial#
		Output_matrices	Est. parameter_col#_row#_trial#
		Output_graphs	Graph_x_y_trial#
PhD BR	Text_docs	Output_tests	Test_parameter_trial#
		Research_prop	RP_trial# & _revised by
		Lit_rev	Topic_subtopic
		Method_desc	Topic_method
		Paper drafts	Paper code_date & revised by
	Budget	Paper_1	Company_date
		Paper_#	Equipment_list
	administration	Price_lists	Budget_date
		Equipment_inv	Item_rec_date (month_year)
		Plan_budget	Item_con_date
Admin	administration	Receipts_dig	
		Contracts_dig	

Organizar,
Nombrar,
Documentar,
Licenciar,
Copias de seguridad



E-ciencia Datos.

Fuente: <https://edatos.consorcioamadrono.es/dataverse/Madrono>

... Fuente: <https://edatos.consorcioamadrono.es/dataverse/Madrono>

Madroño CONSORCIO BETA ciencia Datos

Estadísticas: 249 Descargas

Repositorio de Datos del Consorcio Madroño

Repositorio de Datos UAH, Repositorio de Datos UAM, Repositorio de Datos UC3M, Repositorio de Datos UNED, Repositorio de Datos UPM, Repositorio de Datos URJC

Buscar en este dataverse... **Buscar** Búsqueda Avanzada

1 a 10 de 17 Resultados

Dataverses (6)
Datasets (11)
Ficheros (85)

Categoría del Dataverse
Organización o Institución (6)

Fecha de Publicación
2015 (8)
2016 (7)
2017 (2)

A Python program to reduce electrical networks
03-abr-2017 • Dataverse Repositorio de Datos UC3M
Ledesma Larrea, Pablo, 2017, "A Python program to reduce electrical networks", doi:10.21950/UWEE7A, Dataverse e-cienciaDatos, V1
Modern transmission networks may contain tens of thousands of busbars. Joint studies involving transmission and distribution grids pose the problem of large size of the resulting system, which may make them unmanageable. This work presents a program that reduces an electrical net...

An ecosystem services approach to the ecological effects of salvage logging: valuation of seed dispersal
17-mar-2017 • Dataverse Repositorio de Datos UAH

LAS LICENCIAS (OKF)

- **OpenDataCommons Attribution License (ODC-BY)**, solo obliga a atribuir/reconocer el propietario original. Permite la copia, distribución y uso para producir obras derivadas sobre ella y modificar, transformar y construir sobre ella para cualquier propósito.
- **PDDL**, *Public Domain Dedication and License*, similar a la CC más permisiva (CC0) sin atribución ni limitación geográfica de uso.
- **Open Database License, ODbL**, es más restrictiva, pero permite compartir –copiar, distribuir y usar la base de datos—, producir trabajos sobre ella y modificar, transformar y construir otros productos sobre la base de datos. Mantiene la (Attribution y obliga a ceder el nuevo producto con esta misma licencia (Share-Alike). Añade condiciones a las bases de datos derivadas y puede aplicar restricciones tecnológicas (DRM) a los datos originales o los derivados. Por tanto, reduce la interoperabilidad y la reutilización

- ✓ **Ciencia abierta: definiciones y contexto**
- ✓ **La importancia de abrir datos**
- ✓ **Beneficios de gestionar bien**
- ✓ **¿Qué se pide en los proyectos y en las revistas?**
- ☐ ¿Cuál es la realidad actual?
- ☐ Conclusiones y debate

¿QUÉ SE PIDE EN LOS PROYECTOS?

H2020



Cómo cumplir con los mandatos sobre gestión y publicación de datos en Horizonte 2020

Programa Horizonte 2020 (art. 29.3)

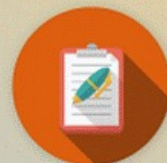


A quién afecta

Investigadores con proyectos subvencionados por Horizonte 2020



Por razones de confidencialidad, seguridad, explotación industrial...
puede no publicar sus datos



Qué obliga a depositar

Los datos, incluidos sus metadatos, necesarios para validar los resultados presentados en las publicaciones científicas

Otros datos, incluidos sus metadatos, especificados en los planes de gestión de datos de los proyectos de investigación



Requisitos

Desarrollar y mantener un Plan de Gestión de Datos

Depositar los datos en un repositorio de datos de investigación

Indicar qué herramientas se requieren para usar los datos

Permitir el acceso, explotación y disseminación de datos

Ventajas

- Permiten **validar los resultados** presentados en publicaciones científicas y otras fuentes de información
- Permiten basarse en los **resultados** de investigaciones previas
- **Fomenta la colaboración** y evita la duplicación de esfuerzos
 - **Acelera la innovación**
- Mejora la **transparencia** del proceso científico

Recomendado el uso de licencias Creative Commons



Open // As open as possible, as closed as necessary // Gestión desde el diseño

4. Los resultados de la investigación de las actuaciones financiadas, incluyendo tanto los resultados difundidos a través de publicaciones científicas como los **datos generados** en la investigación, deberán estar disponibles en acceso abierto, con las excepciones señaladas en el último párrafo de este apartado.

Las publicaciones se depositarán en repositorios institucionales o temáticos de acceso abierto en un plazo que se determinará en las convocatorias en función del área de investigación, que será entre seis y doce meses tras su publicación en las revistas académicas.

Las convocatorias podrán prever que los **datos generados** sean depositados en un repositorio de datos de investigación en abierto, siguiendo un plan de gestión de datos que formará parte de la documentación de la solicitud y que podrá ser modificado durante el periodo de ejecución de la actuación previa información y aprobación del órgano concedente. En la solicitud de la ayuda se indicará si procede o no la difusión en abierto de los resultados, así como de los datos que se generen como resultado de la investigación. En caso afirmativo, se indicará además el repositorio institucional o temático de acceso abierto en el que se depositarán. En caso negativo, se señalará la concurrencia de alguna de las excepciones indicadas a continuación.

Las excepciones a la obligatoriedad del régimen de publicación en abierto y del acceso abierto a los datos de investigación en el caso de que la convocatoria lo prevea, se producirán en los siguientes casos:

a) Cuando se prevea que los **datos generados** en la investigación y los resultados de la investigación realizada puedan ser sometidos a solicitud de la protección de derechos de propiedad industrial o intelectual.

La titularidad de los indicados derechos de propiedad industrial e intelectual se determinará conforme a la normativa específica que les resulte de aplicación.

b) Cuando por su naturaleza los datos estén sujetos a la protección de datos de carácter personal o cuando afecten a la seguridad pública.

"A goal without a plan is just a wish." Antoine de Saint-Exupery (1900 -1944)

Describe el ciclo de vida de la gestión de datos para los datos que deben ser recogidos, procesados y/o generados por un proyecto Horizonte 2020. Como parte de la tarea de hacer que los datos de investigación sean localizables, accesibles, interoperables y reutilizables (FAIR), debe incluir información sobre:

- el tratamiento de los datos de la investigación durante y después de la finalización del proyecto
- qué datos se recogerán, procesarán y/o generarán
- qué metodología y estándares se aplicarán
- si los datos serán compartidos o de acceso abierto y
- cómo se curarán y conservarán los datos (incluso después del final del proyecto).

FAQs <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/support/faq;categories=;programme=null;actions=;keyword=ordH2020>
https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

¿QUÉ SE PIDE EN LOS PROYECTOS?



10 pasos para elaborar un Plan de Gestión de Datos

Un **Plan de Gestión de Datos** (PGD) o Data Management Plan (DMP) es un **documento formal, que debe presentarse al inicio de la investigación, en el que se describe qué**

vas a hacer con tus datos durante y después de finalizar tu investigación y que puede modificarse si se producen cambios en el proceso de la misma.

¿Por qué?

Es una **buena práctica**, es un **elemento clave de Open Science** y es **obligatorio** en los nuevos proyectos H2020.

Herramientas gratuitas para elaborar un PGD



PGDonline
(Consorcio Madroño)
<http://dmp.consociomadroño.es/>



DMPonline (Digital Curation
Centre, UK)
<https://dmponline.dcc.ac.uk/>



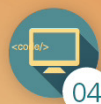
Revisa los **requerimientos** de la entidad financiadora (H2020).



Identifica los **datos**: tipología, procedencia, volumen, formatos y ficheros.



Define cómo se **organizarán y gestionarán los datos**: nombre de los ficheros, control de versiones, software necesario...



Explica cómo se **documentarán los datos**: identifica la información a procesar, consulta si hay estándares o esquemas de metadatos, identifica herramientas que permitan gestionarlos.



Describe los procesos que aseguran una **buena calidad de los datos**.



Prepara una **estrategia de almacenamiento** (durante el proceso) y de preservación de datos (repositorio).



Define las **políticas de datos del proyecto**: cuestiones sobre propiedad intelectual y cómo se tratarán los datos sensibles y personales.



Describe cómo se **difundirán los datos**: dónde, cuáles, cuándo se van a difundir. Si publicarás los datos en un repositorio, como información suplementaria del artículo o como un "data paper".



Asigna **roles y responsabilidades** para las personas y organizaciones participantes en el proyecto.



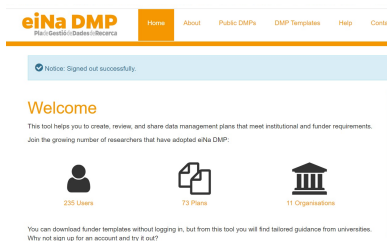
Prepara un **presupuesto realista**: la gestión de datos cuesta tiempo y dinero en términos de software, hardware, servicios y personal.



El personal de tu Biblioteca te puede asesorar



PGD=DMP



<https://dmp.csuc.cat/>



- **Referencia y nombre del conjunto de datos**
Se creará un identificador para el conjunto de datos.
- **Descripción del conjunto de datos**
Descripción de los datos que serán generados o recopilados, su origen (en caso de que sean recogidos), naturaleza y escala, a quién podrían resultar útiles, si sustentan una publicación científica. Información de si existen (o no) datos similares y las posibilidades de su integración y reutilización.
- **Estándares y metadatos**
Referencia a los estándares existentes para la disciplina. Si no existiesen, un esquema sobre cómo y qué metadatos se crearán.
- **Datos compartidos**
Descripción sobre cómo se compartirán los datos, incluyendo procedimientos de acceso, períodos de embargo en su caso, sumarios de los mecanismos técnicos para su disseminación y software necesario y otras herramientas que permitan su reutilización. Definición de si el acceso será abierto o restringido a determinados grupos. Identificación del repositorio en el que se archivarán los datos. Si ya existen y están identificados, indicación del tipo de repositorio (institucional, temático, etc.) En caso de que los datos no se puedan compartir, han de mencionarse las razones (por ej. éticas, jurídicas, protección de datos de carácter personal, propiedad intelectual, comercial, privacidad y seguridad...).
- **Archivo y preservación (incluido el almacenamiento y copias de seguridad)**
Descripción de los procedimientos que se establecerán para la preservación a largo plazo de los datos. Indicación del tiempo que deberían ser preservados los datos, cuál es su volumen final aproximado, qué costes asociados conlleva y como éstos se planificarán y cubrirán.

¿QUÉ SE PIDE EN LAS REVISTAS?

Upload Files

Provide Additional Information

Review & Submit

Guide for Authors

Click Upload Files to select and upload submission files.

- Select a File Type for each submission file.
- Mandatory File Types are indicated in the drop-down list.
- The total size of your submission files may not exceed 700MB.
- The Manuscript File size may not exceed 150MB.
- Update the File Order if necessary, then click Save to preserve the new order before proceeding.
- Please note that in order to proceed with your submission (Only for Review Articles and Research Papers) you are required to upload the author contributions.**
- Click Save & Continue to proceed to the next step.
- Data in Brief:** Authors are invited to submit a Data in Brief article alongside their research article. The Data in Brief article is an attractive alternative, which is normally buried in supplementary material, is actively reviewed, curated, formatted, indexed, given a DOI and publicly available to new, Open Access journal, Data in Brief. Please use the Data in Brief template available through the below link: [Data in Brief Template](#).
- MethodsX (optional)
We invite you to submit a method article alongside your research article. This is an opportunity to get full credit for the time and money you have spent on your method article will be automatically transferred over to the open access journal, MethodsX, where it will be editorially reviewed and published as a separate available [here](#) when preparing your article. Open access fees apply. For any questions, please send your inquiries to mexjm@elsevier.com

CRediT Author Statement
We encourage you to submit an author statement file outlining all authors' individual contributions, using the relevant CRediT roles: Conceptualization; Resources; Software; Supervision; Validation; Visualization; Roles/Writing – original draft; Writing – review & editing. Please format with author name

Data Visualization: Include interactive data visualizations in your publication and let your readers interact and engage more closely with your research with your article.

Enrich your article to present your research with maximum impact. This journal supports the following Content Innovations:

- [Interactive Plots](#): Interactive plot viewer providing easy access to the data behind plots. Please prepare a .CSV file with your plot data and test

Upload Files

Order Files

* File Order	* File Name	* File Type
No data to display.		

Order Files

Save

Share your research data (Optional)

In this section, you can make your research data available with your article. Sharing research data helps other researchers evaluate your findings, and increases the transparency of your protocols, methods and materials.
Please cite your research data within your article; consult your journal's [Guide for Authors](#) for instructions.
The following options are available:

Link research data

If your research data is already hosted in a repository, you can link it to your article here. [Learn more](#)



OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

An author keyword analysis for mapping

Luis-Millán González, Xavier García-Massó, Alberto Pardo-Ibañez, Fernanda

Published: August 1, 2018 • <https://doi.org/10.1371/journal.pone.0201435>

Article	Authors	Metrics
⌵		

Abstract

Introduction

Materials and methods

Results

Discussion

Conclusions

Supporting information

Abstract

Scientific production has increased exponentially in the last thirty years. Likewise, the number of methodological strategies for understanding trends developed in specific fields. Data mining, in particular, our study presents a global analysis in the Sport Sciences Category (SSC) including the frequency of appearance and the dynamics for the last thirty years. Likewise, the nature

VÍAS PARA HACER ABIERTOS LOS DATOS

- Vía dorada del material suplementario como PLOS o ELSEVIER
- Vía verde de los repositorios interoperables



International registry on research data
<http://odisea.ciepi.org>



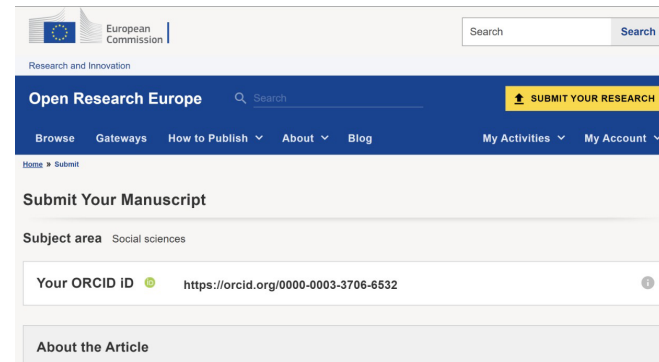
ZENODO

- Desarrollado como repositorio huérfano por el CERN sobre Invenio. Recientemente ha sido relanzado para todo tipo de datos dentro el proyecto europeo OpenAIREplus y probado con datasets del *Wellcome Trust* (Príncipe, 2013).

- Cabe destacar varios aspectos en diferentes momentos de la gestión de datos: verifica que los ficheros estén relacionados con una investigación; asegura un control descentralizado de los datos; y por último permite crear comunidades de usuarios finales.

OPEN RESEARCH EUROPE

25/3/21



European Commission | Research and Innovation

Open Research Europe

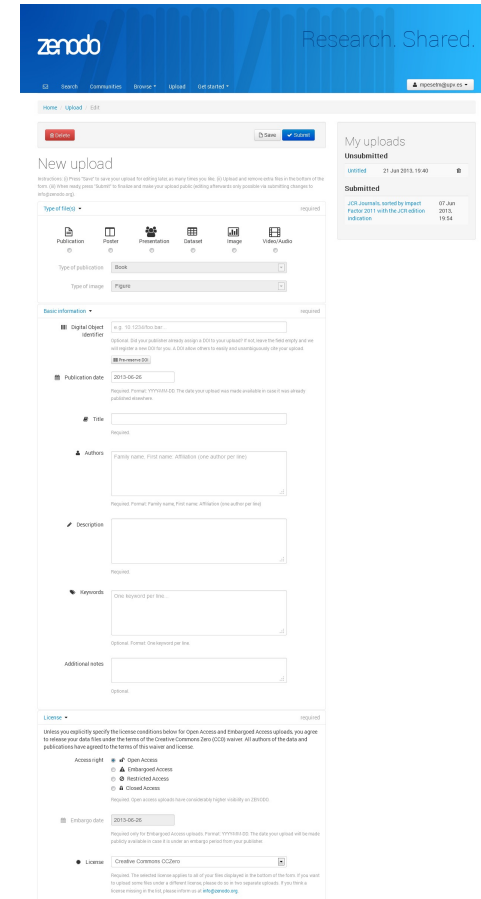
Submit Your Manuscript

Subject area: Social sciences

Your ORCID iD: <https://orcid.org/0000-0003-3706-6532>

About the Article

<https://open-research-europe.ec.europa.eu/author/version/14520>



zenodo | Research. Shared.

New upload

Type of file: Book

Basic information

Input/Client identifier: 10.1111/1111/1111

Publication date: 2013-01-01

Title:

Authors:

Description:

Keywords:

Additional notes:

License:

Embargo date:

[HTTPS://ZENODO.ORG/](https://zenodo.org/)

https://recercat.cat/bitstream/handle/2072/284974/1705_RecomanacionsSeleccionarRepositoriDades.pdf?sequence=5

Nina-Alcocer, Victor; Blasco-Gil, Yolanda; Peset, Fernanda (2013). Datasharing: guía práctica para compartir datos de investigación. El profesional de la información, noviembre-diciembre, v. 22, n. 6, pp. 562-568.
<http://eprints.rclis.org/20907/1/datasharing.pdf>

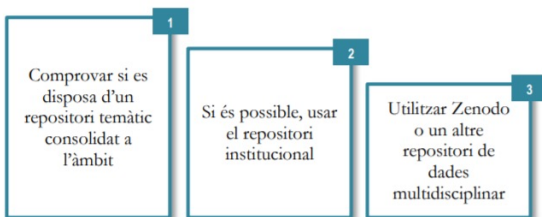


Figura 1. Proposta basada en les recomanacions d'OpenAIRE (vegeu bibliografia)

CSUC







OBSERVATORI VALENCIÀ DE DADES OBERTES I TRANSPARÈNCIA

DATAUSE

m

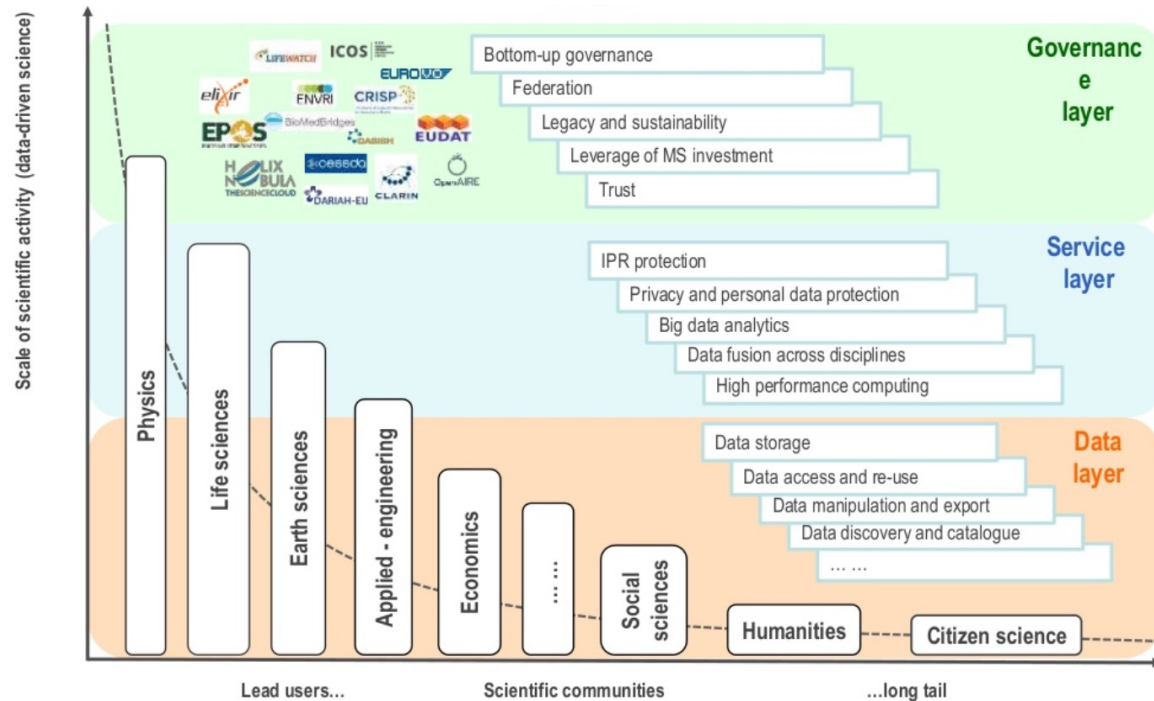
MANO MATHEMATICS & APPLICATIONS

IMI

	 CERN	 EUDAT	 Digital Science	 NCSU	 Harvard University	 Elsevier
1. Àmbit temàtic i geogràfic	Multidisciplinar Internacional	Multidisciplinar Internacional	Multidisciplinar Internacional	Ciència i medicina Internacional	Multidisciplinar Internacional	Multidisciplinar Internacional
2. Tipologia de dades	Tot tipus de formats i també documents publicats	Tot tipus de formats	Tot tipus de formats	Preferència per formats de codi obert, els recomanats	Tot tipus de formats amb funcionalitats addicionals per alguns	Tot tipus de formats
3. Mida aproximada dels fitxers a incloure	Fins a 50 GB per fitxer. Per a mides superiors, contactar amb ells	Fins a 10 GB per fitxer (màxim de 20 GB per registre). Per a mides superiors, contactar amb ells	Fins a 5 GB per fitxer.	300 GB per un conjunt de dades. Per a mides superiors, contactar amb ells	Fins a 2 GB	Fins 10 GB per conjunt de dades. Si hi ha subscripció institucional es poden crear conjunts de dades fins a 100 GB.
4. Dades obertes, embargades, restringides o tancades	- Obertes - Embargades - Restringides - Tancades	- Obertes - Embargades - Restringides	- Obertes - Embargades - Restringides - Tancades (per defecte fins que no es publiquen)	- Obertes - Embargades (si l'editor ho permet, l'autor ho sol·licita en el moment d'enviar les dades a help@datadryad.org) - Restringides (sempre que l'autor ho requereixi mentre es revisa per parells)	- Obertes - Restringides	- Obertes - Tancades (per defecte fins que es validen els requisits) - Restringides (Per institucions) - Embargades
5. Llicències amb les que es volen difondre les dades	Permet varietat de llicències encara que recomanen llicències obertes. La descripció de les dades està sota CC0 .	Permet varietat de llicències per a dades i per a software.	- CC0 per a dades i metadades - CC-BY per a no dades - MIT per a codi i software - Altres llicències: GNU, Apache 2.0, BSD 3 clause.	CC0	Per defecte CC0 però es poden especificar altres termes d'ús .	Permet varietat de llicències per a dades (CC0, CC-BY, CC-BY-NC), software (MIT, Apache, BSD, GPL, LGPL, MPL, CeCILL) i maquinari (CERN, TAPR).
6. Identificador permanent	DOI	Handle i DOI	DOI	DOI	Handle i DOI	DOI
7. Condicions per retirar contingut	- Es permet retirar les dades justificant-ne el perquè. - El DOI i la URL quedaran retinguts.	Sense informació.	Es pot sol·licitar la retirada de contingut públic justificant-ne el motiu.	- Cessió de dades irrevocable . - Es reserva el dret a eliminar el contingut justificant-ne el motiu.	Es permet retirar les dades, tot i que les metadades quedaran sempre visibles.	Es permet la retirada de dades . Per retirar el contingut contactar amb Mendeley .

CSUC (2020). Recomanacions per seleccionar un repositori per al dipòsit de dades de recerca: Versió 5, oct 2020 hdl.handle.net/2072/377586

LA FUTURA EUROPEAN OPEN SCIENCE CLOUD



(Burgelman)

Strategic Research and Innovation Agenda (SRIA)

of the European Open Science Cloud (EOSC)

Version 0.8 18 October 2020



**EUROPEAN OPEN
SCIENCE CLOUD**

<https://www.eoscsecretariat.eu/sites/default/files/eosc-sria-v08.pdf>

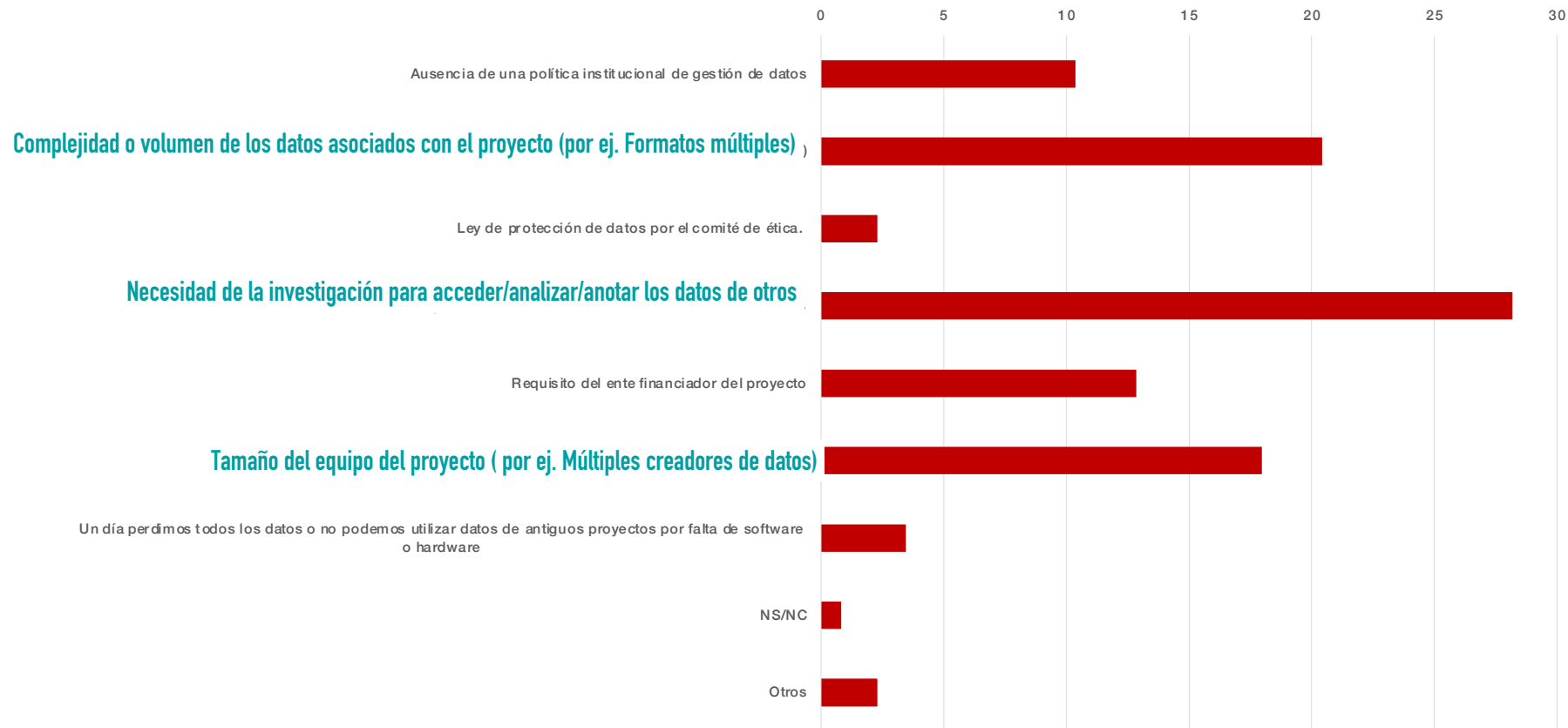
- ✓ **Ciencia abierta: definiciones y contexto**
- ✓ **La importancia de abrir datos**
- ✓ **Beneficios de gestionar bien**
- ✓ **¿Qué se pide en los proyectos y en las revistas?**
- ✓ **¿Cuál es la realidad actual?**
- ☐ **Conclusiones y debate**

NUESTRA REALIDAD

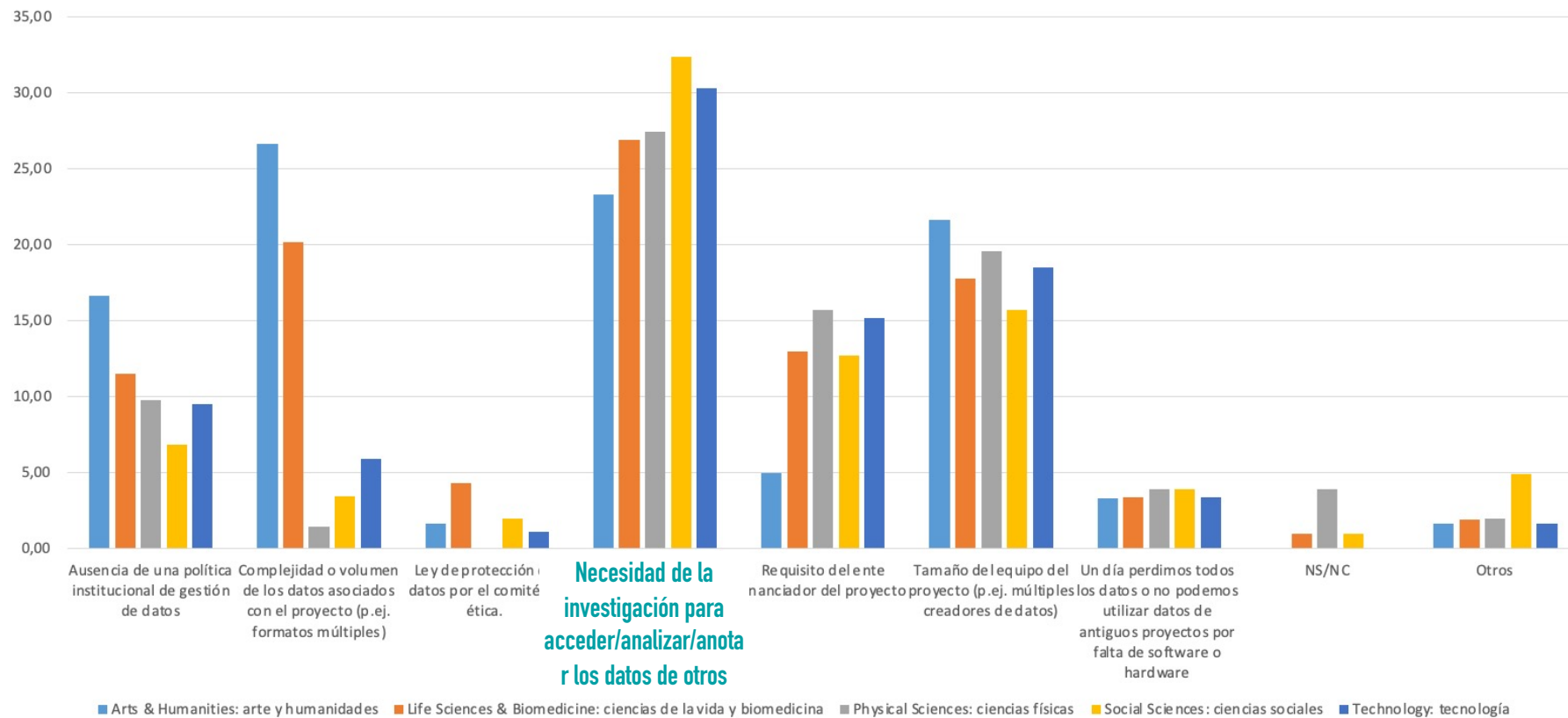
La opinión y los hábitos de los investigadores sobre los datos de investigación.

Aleixandre-Benavent, Rafael; Vidal-Infer, Antonio; Alonso-Arroyo, Adolfo; Ferrer Sapena, Antonia; Peset, Fernanda (2020).
Research data sharing in Spain: exploring determinants, practices and perceptions. *Data* 2020, 5(2),
29; <https://doi.org/10.3390/data5020029> <https://www.mdpi.com/2306-5729/5/2/29>

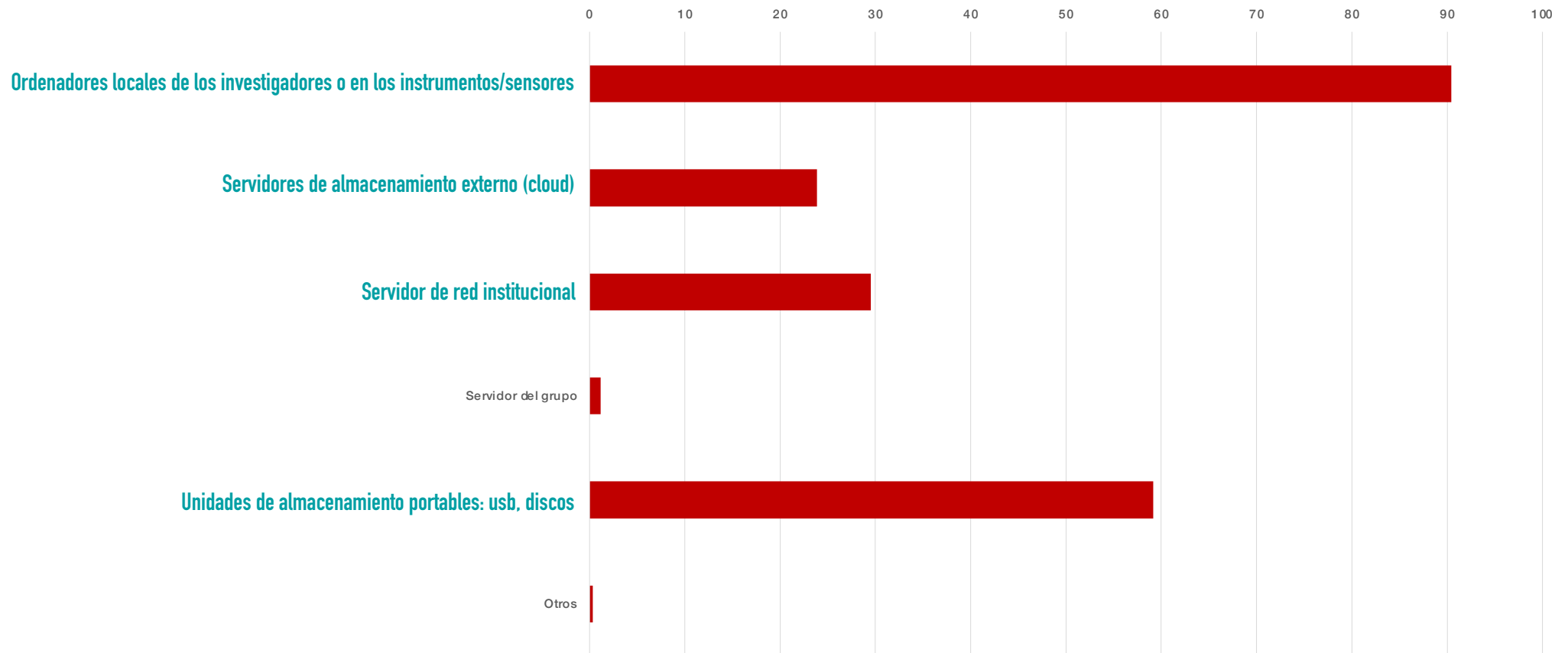
Motivos por los que se desarrolla una política de gestión de los datos



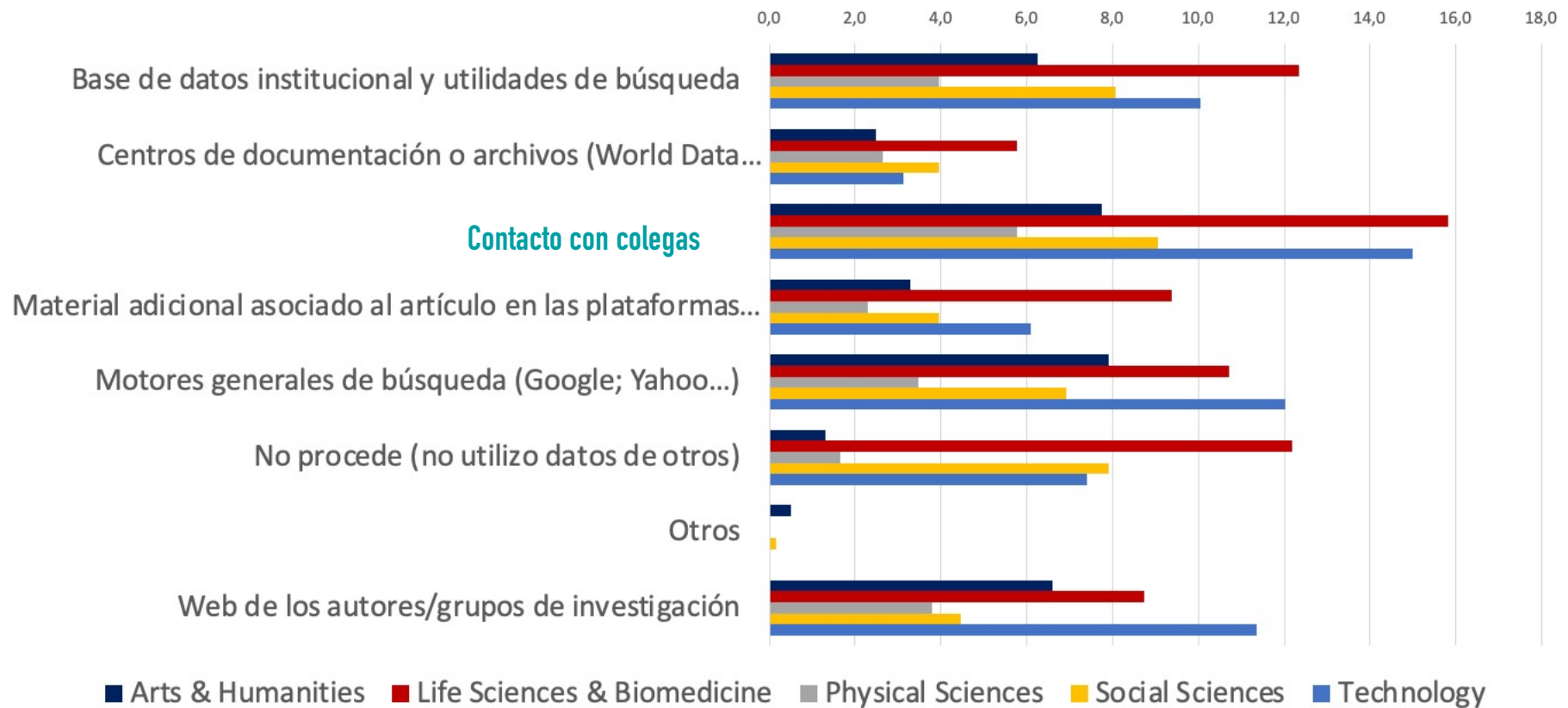
Cual fue el motivo principal para desarrollar la política de datos



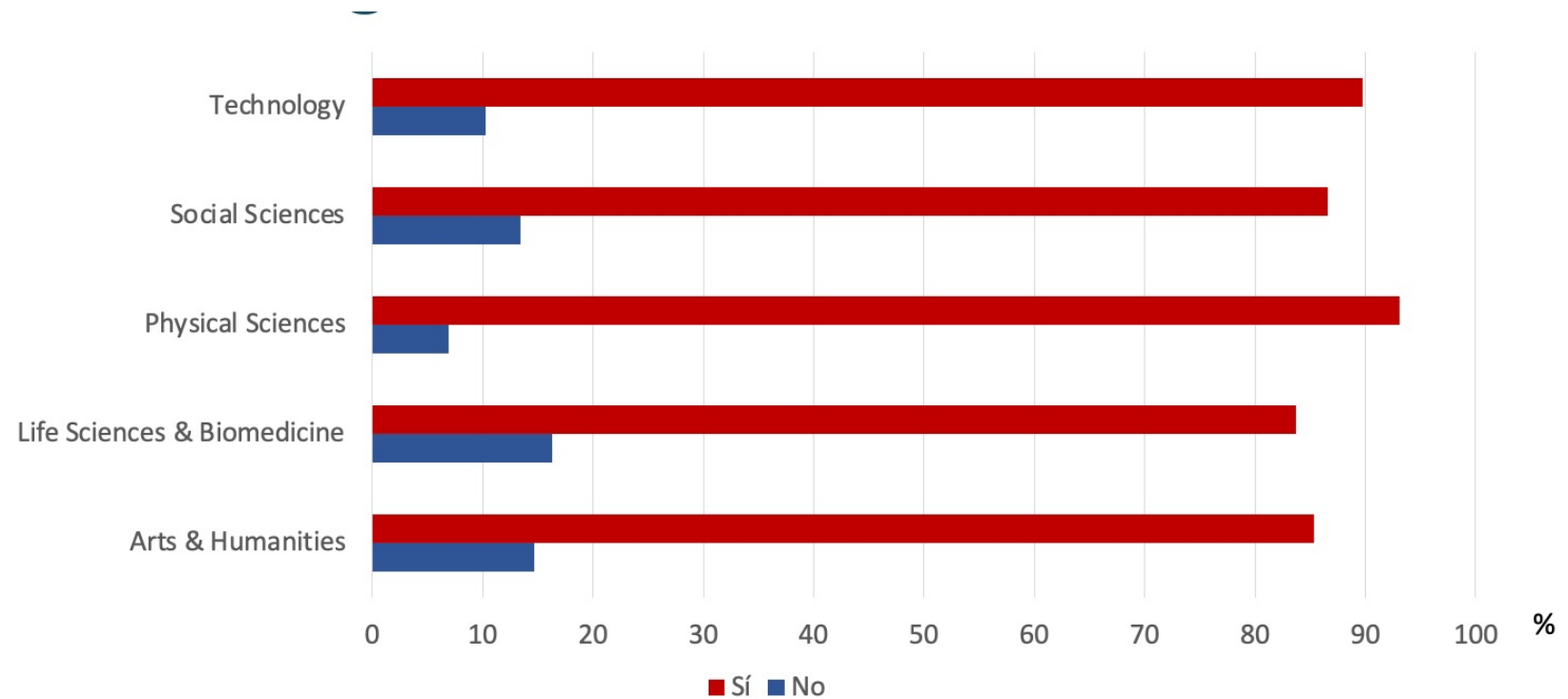
Donde suele almacenar los datos de investigación



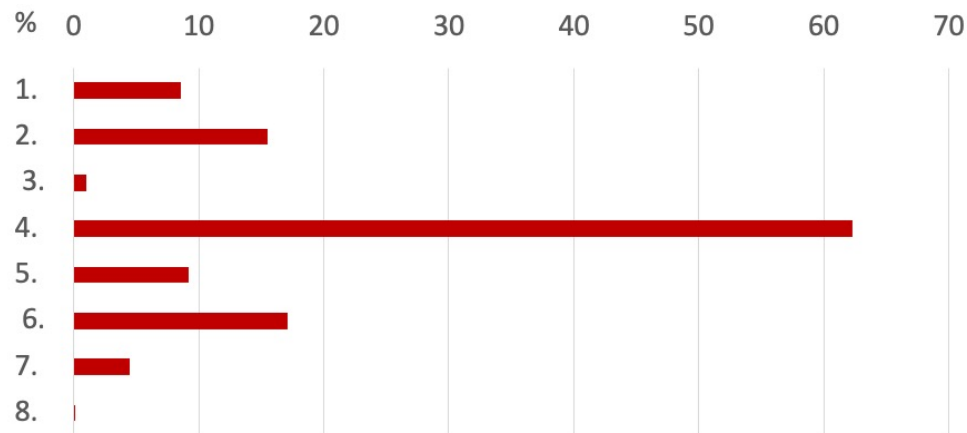
Medios utilizados para localizar y acceder a los datos de otros investigadores



¿Le gustaría utilizar datos de investigación de otros investigadores?



En relación a la disponibilidad de los datos



1. Mis datos están disponibles en abierto para mi disciplina científica

2. Mis datos están disponibles en abierto para todo el mundo

3. Mis datos están disponibles mediante el pago de una tasa

4. Mis datos están disponibles para mi grupo de investigación y/o colegas colaboradores en la investigación

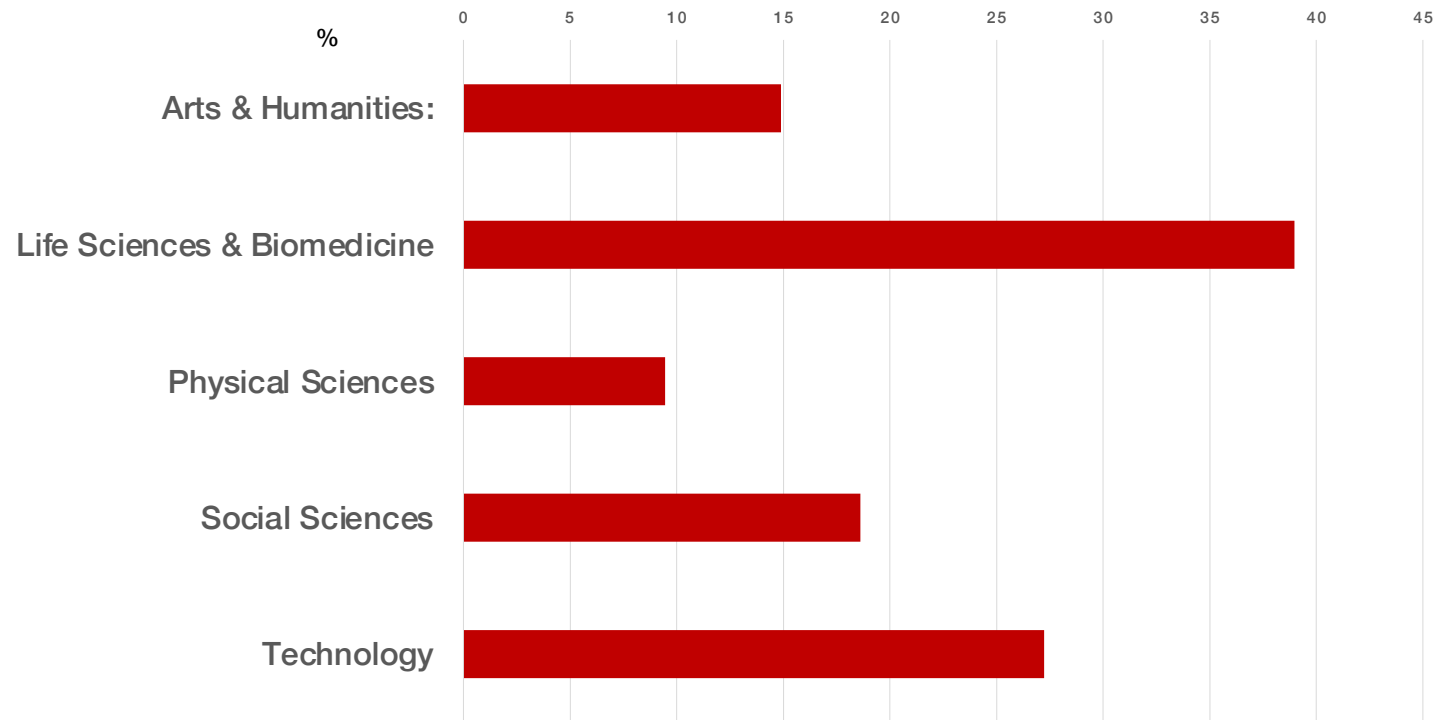
5. Mis datos podrían estar disponibles con los cambios apropiados (por ejemplo: datos clínicos anónimos o con la obligación de citarlos)

6. No comparto mis datos pero me gustaría hacerlo en un futuro

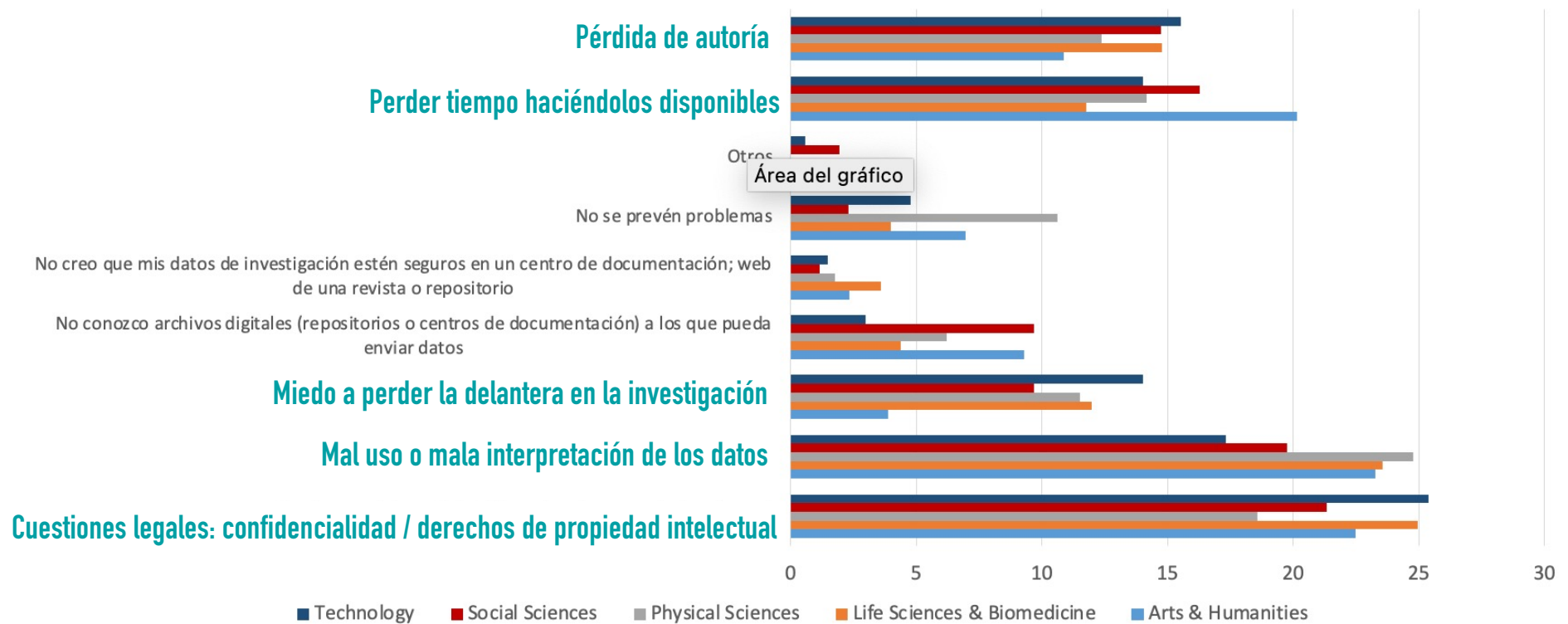
7. No comparto mis datos y no quiero hacerlo en el futuro

8. Otros

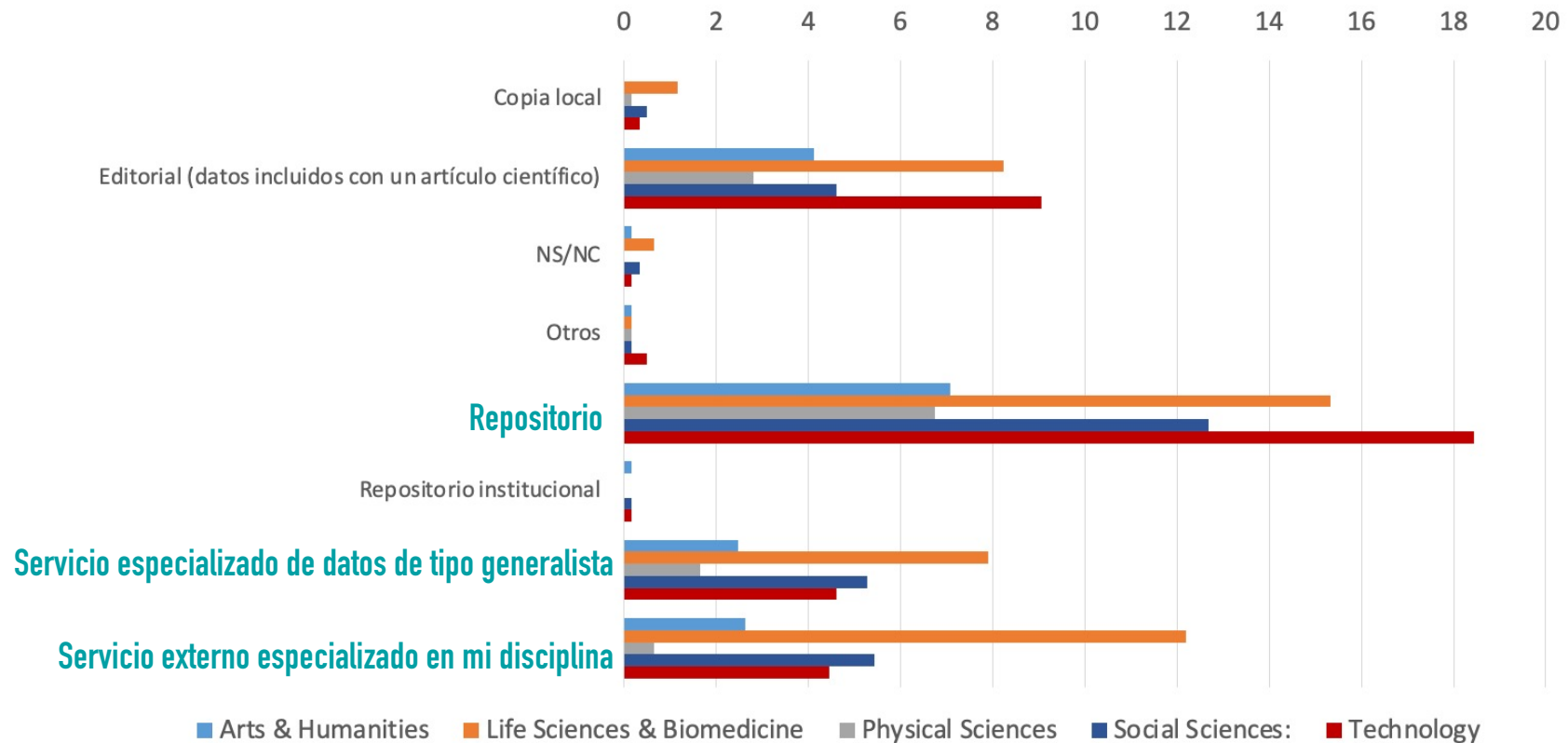
¿Alguna vez le han solicitado sus datos otros investigadores?



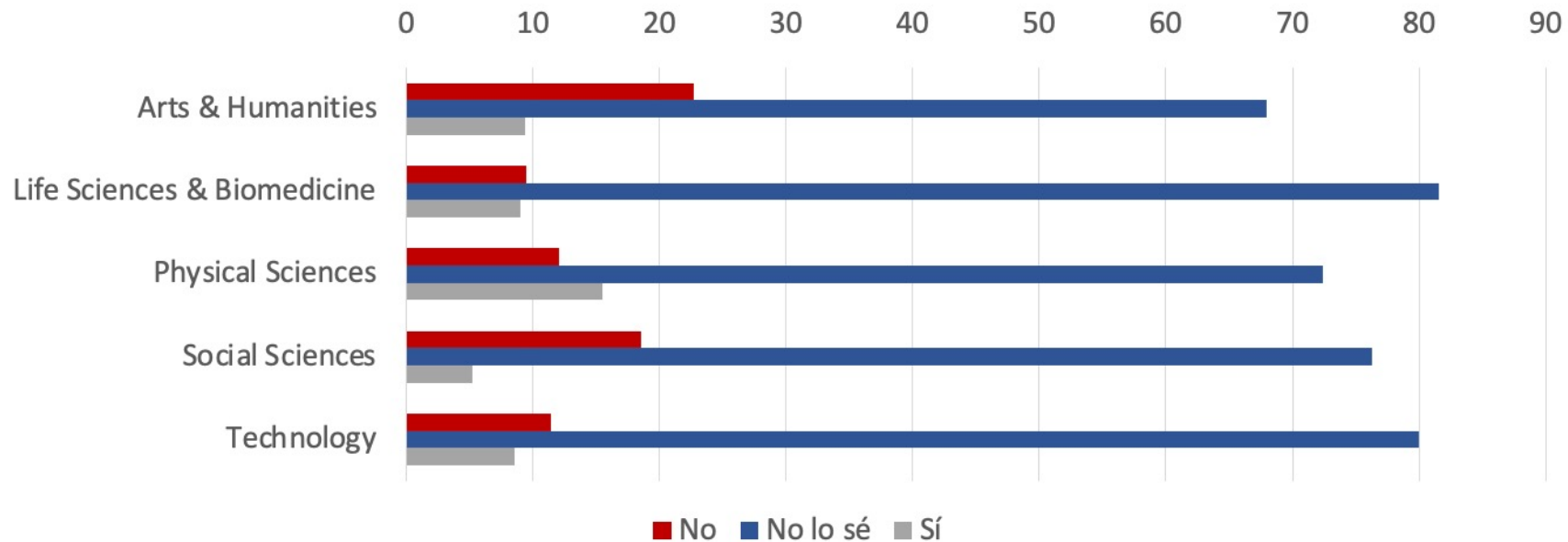
Qué le preocupa a la hora de compartir los datos en un servicio externo



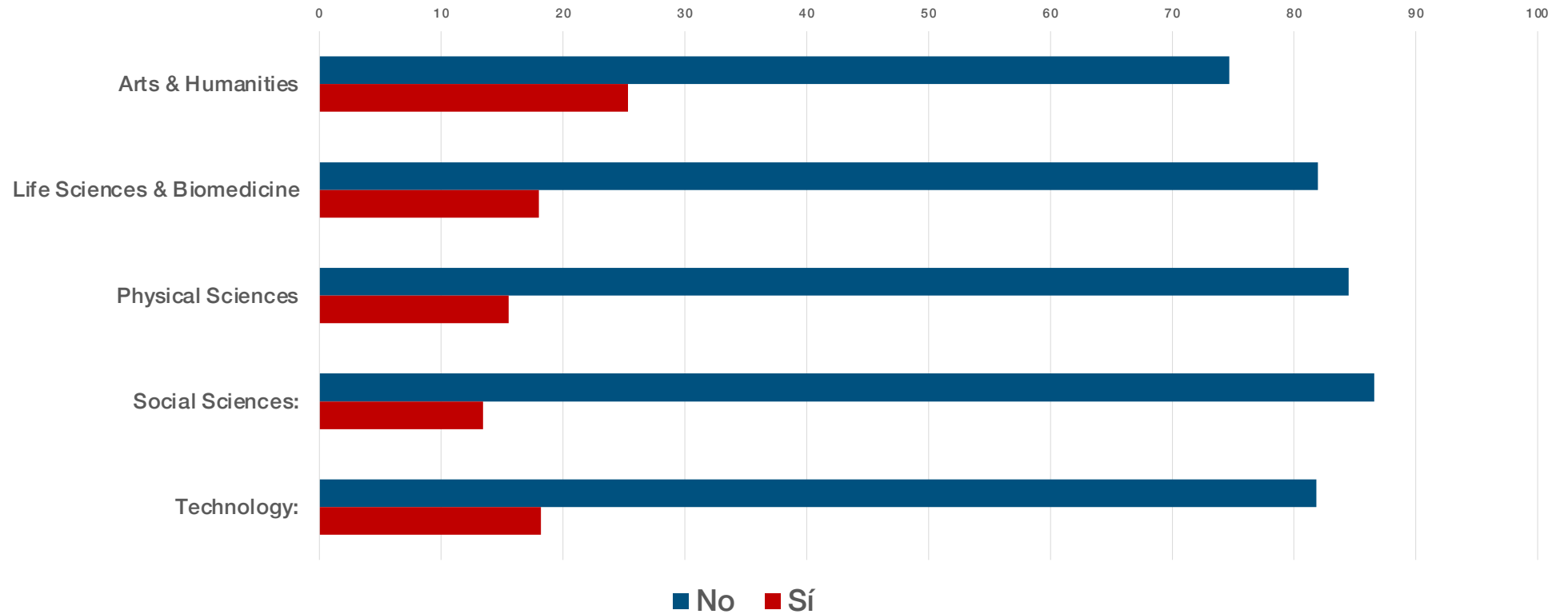
Donde se considera más adecuado la preservación y compartir los datos



Existe en su disciplina alguna instalación de datos digitales que pueda utilizar



Su organización le ayuda con la preservación de datos



Conclusiones en datos de investigación



- Suelen utilizar ordenadores locales o memorias digitales portátiles como sistemas de almacenamiento y depósito.
- A una amplia mayoría les gustaría utilizar los datos de investigación recogidos por otros investigadores.
- La principal amenaza a sus datos son es la restricción de acceso y uso relacionada con la propiedad intelectual.
- El principal motivo para desarrollar una política de datos fue por la necesidad de la propia investigación en todas las áreas excepto en Arte y Humanidades, área en la que la complejidad o el volumen de datos asociados con el proyecto fue el motivo principal.

- ✓ **Ciencia abierta: definiciones y contexto**
- ✓ **La importancia de abrir datos**
- ✓ **Beneficios de gestionar bien**
- ✓ **¿Qué se pide en los proyectos y en las revistas?**
- ✓ **¿Cuál es la realidad actual?**
- ✓ **Conclusiones y debate**

Conclusiones datos abiertos

- Normalización de autoridades
- Definición de conceptos homogénea
- Bases de datos accesibles
- Estructuras de datos normalizadas
- Principios FAIR
- Cinco estrellas de Tim Berners-Lee

IMPACTO POSITIVO

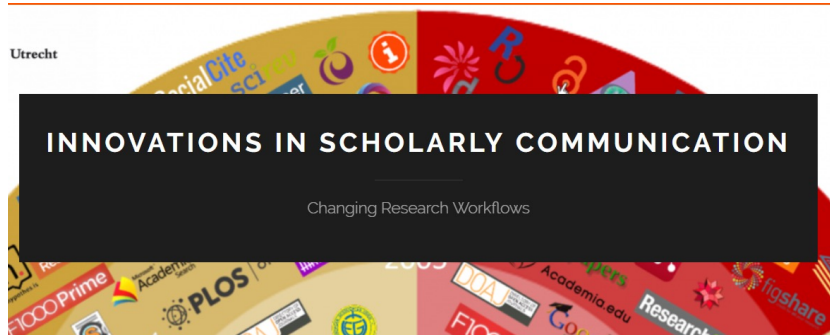


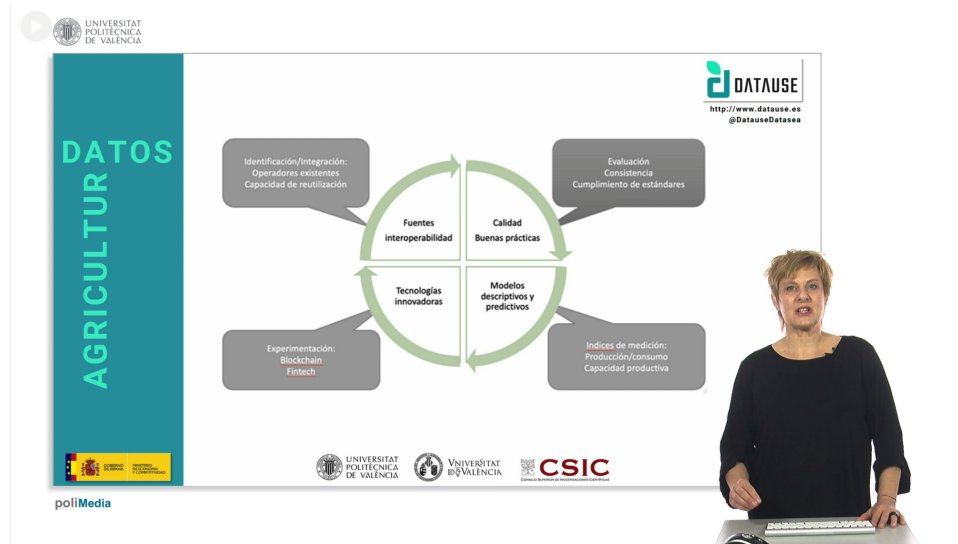
- **Reduce la duplicación en la investigación**, en términos de tiempo, esfuerzo y financiación;
- **Reduce la gestión y administración de los recursos digitales** de los requisitos solicitados por los organismos de financiación.
- **Se amplían los resultados de la investigación** basados en datos existentes integrados y analizados de múltiples disciplinas y regiones.

Acciones necesarias:

- Campañas de sensibilización / Infraestructuras socio técnicas / Definir las reglas
- Todo ello, impulsado por los organismos oficiales

Kramer, Bianca; Bosman, Jeroen (2015):
101 Innovations in Scholarly
Communication - the Changing
Research Workflow. figshare. Poster.
<https://doi.org/10.6084/m9.figshare.1286826.v1>





DATASEA <http://www.datasea.es>
MAREDATA <http://www.maredata.net>
DATAUSE <http://www.datause.es>

Peset, F. (2014). Datause. [@datausedatasea](https://media.upv.es/#/portal/video/f061ca60-69e5-11eb-9339-cbab4b322e5a)



<https://www.ctranspa.webs.upv.es/encuesta-datos-abiertos/>

https://es.surveymonkey.com/r/datos_abiertos_investigadores

¡Gracias!

¡ESPERAMOS SUS PREGUNTAS!